
SERVER BASED OBJECT RECOGNITION

Gopu Sai Manish*¹, Duggineni Rajesh*², Vislavath Dharam Singh*³,

Dr. B. Rajendra Naik*⁴

*^{1,2,3}UG Student, Electronics And Communication Department, University College Of Engineering
Osmania University, Hyderabad, Telangana, India.

*⁴Guide, Electronics And Communication Department, University College Of Engineering
Osmania University, Hyderabad, Telangana, India.

ABSTRACT

In today's modern world real-time object recognition is one of the most intriguing topics and has taken place all over because of its prominent role in automation. However, because of its high complexity in computations and gaps in the methodology, it is highly inefficient and demands the use of high-end processors in autonomous and IoT devices to execute training and recognition. Our project aims to resolve this problem by performing the calculations on a server using the concepts of IoT and YOLO algorithm for training and prediction. The server does all of the intricate calculations and computations without the aid of any external hardware, making image recognition simpler and more efficient.

Keywords: Architecture, YOLO, Bounding-Box, Confidence Score, FIFO.

I. INTRODUCTION

Object recognition is a type of computer vision technology that is used to recognize objects in images. When viewing a picture or a video, we may swiftly identify faces, things, situations, and other visual components. In autonomous and IoT devices, the goal is the same: to comprehend the information of an image. Driverless cars rely on object recognition technology to differentiate between a stop sign and a lamppost. Additionally, it helps a variety of applications, including such robotic vision, fault diagnosis, industrial automation, surveillance, and disease monitoring in bio imaging.

The GSM Module, Camera Module, and Arduino are the key hardware components needed to build this project. The YOLO algorithm and server program put together recognize the final objects in the image, indicating the object and the confidence score. Firstly, the image that we intend to recognize must be captured by the camera. The OV7670 camera module is a better choice because it has a 640×480 VGA output resolution and uses less power, making it economical. Then, the Arduino UNO, which comes with the ATmega328P microcontroller, builds a bridge between the camera module and the GSM module by processing and forwarding image data. TCP/IP protocols are then used by the GSM Module, which is configured using AT commands, to send this data to the server over the internet. After processing this image data in the server using the YOLO algorithm, the results show the presence of objects and their locations.

II. METHODOLOGY

The OV7670 camera module is used to capture the image. Arduino-UNO collects the image data from camera module and is transmitted to the servers using the SIM 900A GSM module via TCP communication between SIM 900A and the server. GSM module is controlled using AT commands sent by Arduino UNO using serial communication. The server then collects the image data and processes it to predict the objects in it using the pre-trained YOLO algorithm. Depending on the application, this result may be either transmitted back to Arduino through the GSM module or stored on a server or in the cloud.

SIM 900A GSM Module:

GSM (Global System for Mobile communications) is a digital cellular technology used for transmitting mobile voice and data services. The GSM/GPRS-compatible Quad-band device can provide data transfer speeds of up to 9.7 Kbits per second, allowing the transmission of basic SMS. It can also be used for various applications such as surfing the Internet, calling, and sending and receiving text. It mainly works through a set of predefined instructions called AT Commands, where AT stands for Attention.

Arduino UNO:

It is a ATmega328 microcontroller board having 20 digital input/output pins, 16MHz resonator, a power jack a USB connection for dumping the code etc. It has 32KB of flash memory for code storing purpose. The Arduino integrated development environment (IDE) is a cross-platform application written in Java, and is derived from the IDE for the Processing programming language.

OV7670 Camera Module:

The camera module is powered from a single +3.3V power supply. An external oscillator provides the clock source for camera module XCLK pin. With proper configuration to the camera internal registers via I2C bus, then the camera supply pixel clock (PCLK) and camera data back to the host with synchronize signal like HREF and VSYNC. The OV7670 camera module is a low cost 0.3 mega pixel CMOS color camera module, it can output 640×480 VGA resolution image at 30fps. The OV7670 camera module build in on board LDO regulator only single 3.3V power needed and can be used in Arduino, STM32, Chip kit, ARM, DSP, FPGA and etc.

The data flow via several components to produce the final output is shown in the following diagram.

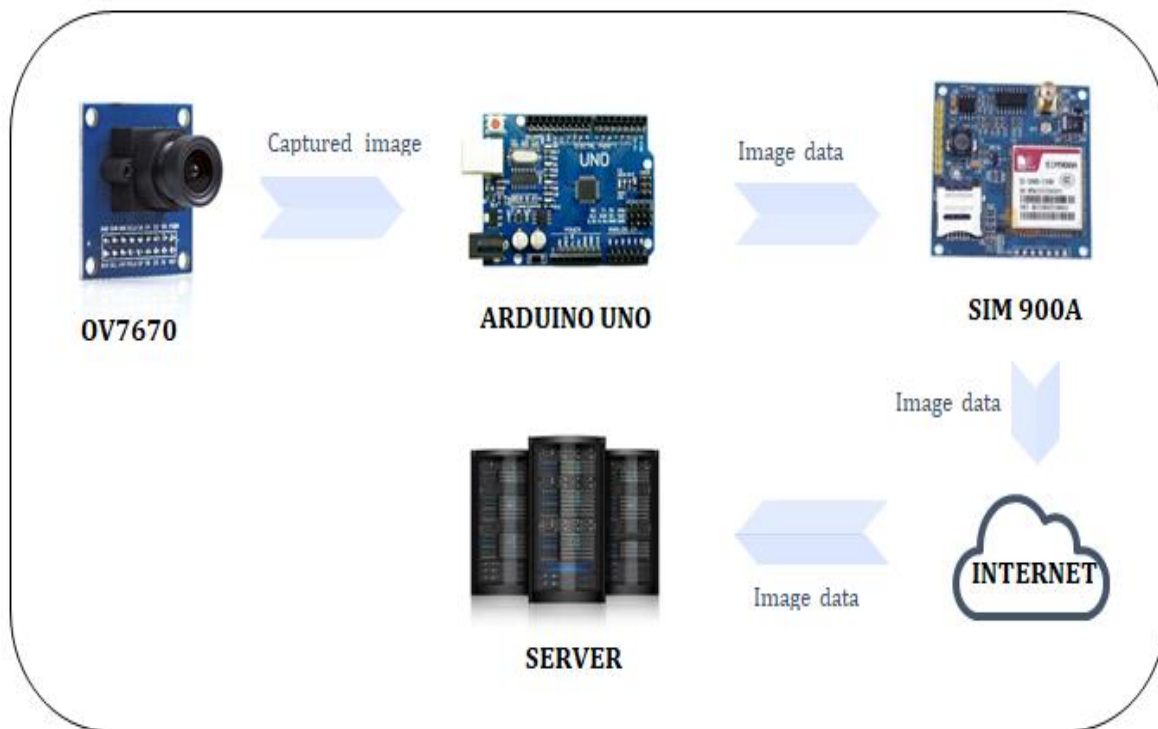


Figure 1: Block Diagram

III. MODELING AND ANALYSIS

OV7670 Camera Module is a FIFO camera Module available from different Manufacturers with different pin Configurations. The OV7670 provides full frame, windowed 8-bit images in a wide range of formats. The image array is capable of operating at up to 30 frames per second (fps) in VGA. The camera module OV7670 can be interfaced with Arduino UNO with same pin configuration, code and steps. The camera module is hard to interface because it has large number of pins and jumbled wiring to carry out. Also, the wire becomes very important when using camera modules as the choice of the wire and length of the wire can significantly affect the picture quality and can bring noise. To initialize the camera, we only have the options to change the register values. The register values need to be changed from the default to the custom. Also, we need to add required delay depending upon the microcontroller frequency we are using.

It is interfaced with Arduino using connections shown below.

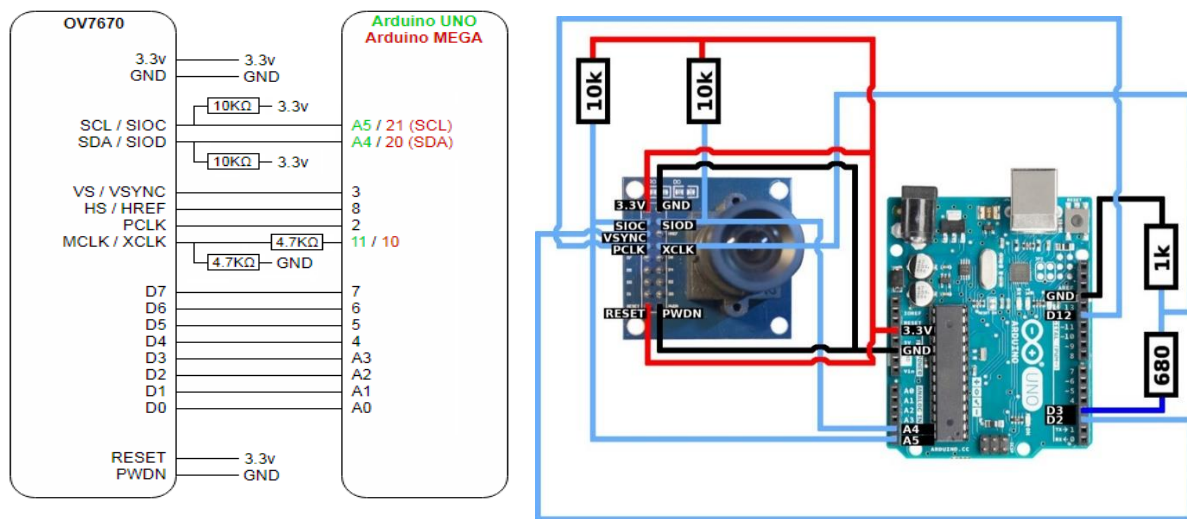


Figure 2 & 3: Pin configuration for interfacing camera module with Arduino

The SIM900A module has 6 pins in which two pins for V_{cc} and Ground and the rest are 3VR & 3VT (3volt Rx & Tx) and 5VR,5VT (5volt Rx & Tx) and the connections are made as follows:

- V_{cc} to 5V
- Ground to Ground
- 5VR digital pin 7
- 5VT digital pin 8

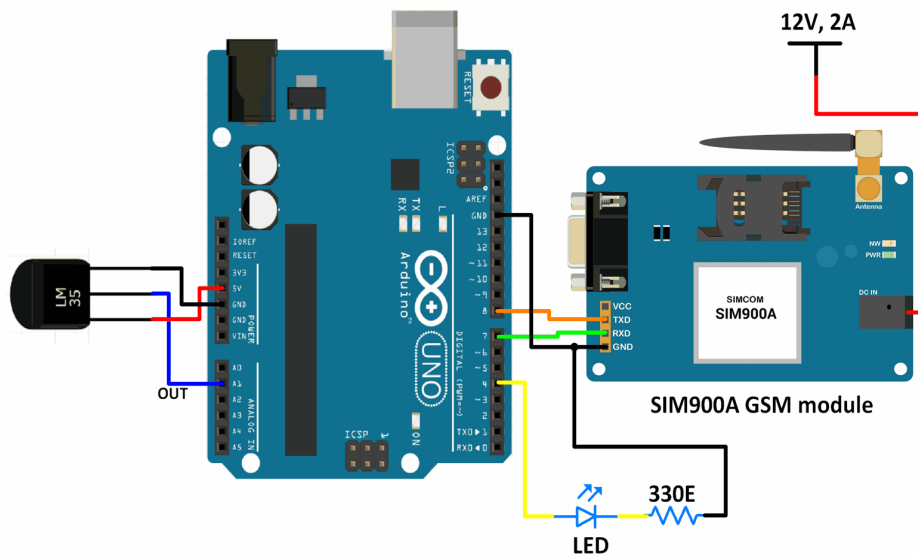


Figure 4: Interfacing GSM module with Arduino

Although object identification is a simple and quick task for humans to do, because of the algorithmic nature of the task, it is relatively challenging to accomplish on a computer. It is possible to characterize the object identification issue as a labeling issue based on models of well-known items. Formally, the system should assign the correct labels to areas, or a group of regions, in the image given comprising one or more objects of interest (and background), as well as a set of labels matching a set of models known to the system. YOLO architecture eliminates this problem by using prior known object models and locates objects in an image of the real world.

IV. RESULTS AND DISCUSSION

Having seen the implementation and advantages of YOLO, the results that are obtained by applying this algorithm to the real time images are analyzed.

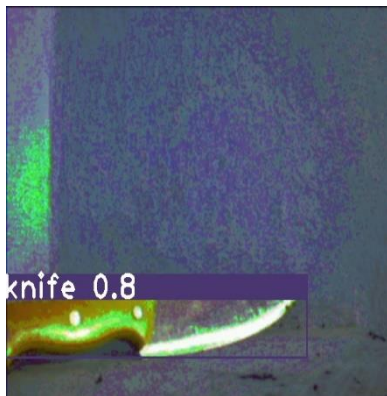


Figure 5: Knife



Figure 6: Scissor

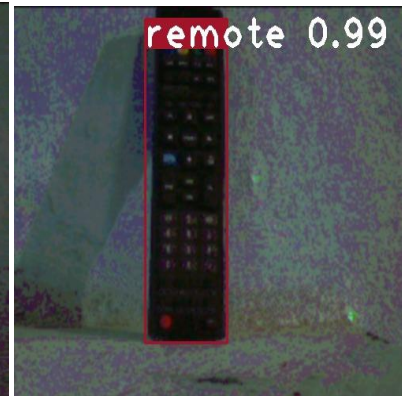


Figure 7: Remote

In the above figures, the algorithm has successfully detected the objects as a knife in Fig. 5, a scissor in Fig. 6, and a remote in Fig. 7, and given the bounding box for each of the objects. In addition to the bounding box, the algorithm also gave a confidence score for that box of 0.8 for the knife, 0.94 for the scissor, and 0.99 for the remote respectively.

Apart from recognizing non-living objects this project promises in identifying living beings as well. Below is a sample of recognizing a human.



Figure 8: Person

In addition to the above where only objects belonging to a single class are present, the algorithm can successfully detect multiple objects belonging to different classes in the image. This is possible only because of the YOLO algorithm since it can predict multiple objects in an image at once. Thus, the output is obtained very quickly. Here are some more results involving multiple objects belonging to different classes.

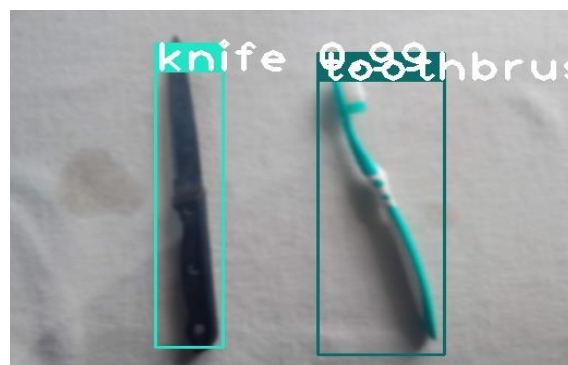


Figure 9: Knife and Tooth brush

V. CONCLUSION

The YOLO algorithm, which makes use of a single neural network, is suggested for the purpose of object recognition. When applied to other domains and generalized from natural images, this methodology performs better than other approaches. The algorithm is simple to implement and may be trained on an entire image. Object recognition is being more widely adopted by businesses, with applications ranging from personal security to workplace productivity. Application areas for object identification in image processing include photo retrieval, security, observation, automated driving systems, and machine analysis.

VI. REFERENCES

- [1] Geethapriya.S, N. Duraimurgan, S.P Chokkalingam, "Real time object detection with YOLO", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019.
- [2] Joseph Redmon, Santosh Divvala, Ross Girshick, "You Only Look Once: Unified, Real-Time Object Detection", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.
- [3] Surya Ramanan, "A beginner's guide to Object Detection Algorithms", <https://medium.com/analytics-vidhya/beginners-guide-to-objectdetection-algorithms-6620fb31c375>, Apr 28 ,2019.
- [4] Amro Kamal, "YOLO, YOLO v2, YOLO v3 all you want to know", https://medium.com/@amrokamal_47691/yolo-yolov2-and-yolov3-allyou-want-to-know-7e3e92dc4899, Aug 03,2019.
- [5] Agarwal, S., Awan, A., and Roth, D. (2004). Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. Pattern Anal. Mach. Intell. 26, 1475–1490. doi: 10.1109/TPAMI.2004.108
- [6] Pulkit Sharma, "A practical to object detection using YOLO framework", <https://www.analyticsvidhya.com/blog/2018/12/practical-guide-objectdetection-yolo-framework-python/>, Dec 6,2018.
- [7] Abdul Vahab, Maruti S Naik, Prasanna G Raikar, Prasad S R, "Applications of Object Detection Systems" International Research Journal of Engineering and Technology (IRJET) ISSN: 2395– 0056, Volume:06, Issue:04, Apr 2019.